知识管理论坛 ISSN 2095-5472 CN11-6036/C
Knowledge Management Forum

E-mail: kmf@mail.las.ac.cn
http://www.kmf.ac.cn

【学术探索】

基于实体关系的企业并购战略意图研究

◎ 廖建军 [1,2,3]

[1] 中国科学院文献情报中心 北京 100190 [2] 中国科学院大学 北京 100190
[3] 郑州航空工业管理学院 郑州 450015

摘要：[目的 / 意义] 针对竞争情报中企业战略情报的获取问题，提出一套自动高效的战略意图挖掘流程，帮助企业更好的收集战略情报信息，提高战略意图挖掘的有效性。［方法 / 过程］引入实体关系抽取技术探索获取企业战略情报的系统方案，以企业并购事件为例，应用关系抽取技术收集企业并购关系，采用实体属性抽取技术揭示企业并购的战略意图，通过阿里巴巴集团的并购事件检验抽取系统的效果。［结果 / 结论］该方案能够有效地从海量网络信息中自动抽取企业并购关系并挖掘出企业经营领域，通过可视化地呈现揭示企业并购的战略意图，达到快速、自动地获取企业竞争情报的目的。

关键词：实体关系　关系抽取　企业并购　战略意图　竞争情报
分类号：G250

引用格式：廖建军．基于实体关系的企业并购战略意图研究 [J/OL]. 知识管理论坛 , 2016, 1(4): 266-275[引用日期]. http://www.kmf.ac.cn/p/1/45/.

1 引言

企业并购目前已然成为中国互联网公司的主要趋势。2015 年 11 月 6 日，阿里巴巴集团和优酷土豆集团共同宣布，双方已就收购优酷土豆股份签署并购协议，阿里巴巴将拿出大约 45 亿美元现金收购优酷土豆。阿里巴巴是一家电子商务公司，优酷土豆是娱乐视频公司，阿里巴巴收购与自己经营性质不同的优酷土豆的战略意图是什么？

互联网上的公众媒体中存在着大量的公开信息，基于公开信息源获取商业竞争情报成为当前的研究热点。在企业竞争情报中，用户最关心的是竞争对手的战略规划情报，但这也是最难获取到的。散布在互联网上的企业并购事件能够在一定程度上反映出企业的战略规划，因此如果能够从纷繁复杂的网页中抽取企业并购关系，将有助于企业更全面地剖析竞争对手的战略意图，达到知己知彼的目的。然而，目前的研究大多关注竞争情报在企业并购过程中的应用 [1-3]，鲜有通过企业并购研究竞争对手的战略意图的研究。因此，从海量的新闻网页中抽取出企业间的并购关系，揭示出竞争对手的战略意图，对于企业竞争战略情报的获取具有十

作者简介：廖建军，中国科学院文献情报中心、中国科学院大学博士研究生，郑州航空工业管理学院讲师，E-mail: liaojianjun@mail.las.ac.cn。
基金项目：本文系国家社会科学基金青年项目“大数据环境下数字出版知识服务研究”（项目编号：14CTQ028）研究成果之一。
收稿日期：2016-02-15　　发表日期：2016-08-15　　本文责任编辑：徐健

知识管理论坛，2016(4):266−275
DOI：10.13266/j.issn.2095−5472.2016.032

分重要的意义。

2 企业并购

企业并购（merger & acquisition）一般是指一家企业用现金或者有价证券购买另一家企业的股票或者资产，以增强企业本身的竞争优势，实现企业经营目标的行为。企业并购有狭义和广义之分，狭义的并购一般包括兼并和收购，而广义的并购除了兼并和收购外，还包括其他各种形式的资产重组。

2.1 企业并购与竞争情报

竞争情报界以研究企业并购著名的美国杨百翰大学的 P.Dishman 博士曾多次公开强调：“企业并购是一项竞争情报活动”[4]。我国著名情报专家包昌火指出：竞争情报是关于竞争环境、竞争对手与竞争策略的信息和研究[5]。竞争情报是一个过程，通过收集商业信息，识别和分析竞争对手的能力、计划、意图、优点和缺点，转化为情报以辅助商业决策。企业并购是企业发展过程中最重要的战略决策之一，而竞争情报对于决策的成败有着至关重要的作用。

企业并购战略属于战略情报范畴，通常是由企业高层制定，需要通过竞争情报手段对被并购企业进行尽职调查，确保企业并购的成功。在做企业并购战略规划的竞争情报调查时，首先要明确自身的并购需求，还要掌握被并购企业的背景与历史、行业与企业的营销方式、制造方式、财务资料、研究与发展计划等各种相关的问题。因此，企业并购也是一项竞争情报活动。竞争情报贯穿于企业并购前、并购中和并购后的所有过程，通过竞争情报分析方法收集、分析、筛选和评估潜在的并购目标，从而为并购战略提供决策支持。

2.2 企业并购的战略意图

企业并购的目的是充分发挥双方的经营协同效应、市场份额效应等优势，增强企业自身的市场竞争力。这也充分体现了企业的发展战略意图。G. Hamel 和 C. K. Prahalad 认为[6]战略意图是：一个雄心勃勃的宏伟梦想，它为企业提供了动力，能够让企业迈上未来的成功之路。简单来说，战略意图是企业长期发展想要实现的愿景。获取竞争对手的情报最常使用的方法是收集竞争对手的战略行动计划，包括产品研发、新业务开发和企业并购等。显然，对于获取竞争对手的企业并购战略来说，竞争对手的企业并购实施就是最好的情报信息源。企业收购关系是指一个企业收购另一个企业而形成的关系。通常，企业收购关系反映了主导企业的某些战略意图，因此对于竞争情报分析至关重要。例如，2014 年 5 月，Google 收购了英国大数据公司 Rangespan，这表明 Google 要通过大数据技术强化其电子商务业务的战略。

企业并购的过程和变化趋势在一定程度上反映了企业战略动态、市场经营倾向等信息。如果一个企业能够获取其竞争对手的企业并购关系信息，对了解竞争对手的经营发展战略，达到知己知彼以增强其核心竞争力无疑大有裨益。本文依据企业并购的经营协同效应、市场份额效应，提出了一种新的基于实体关系抽取的方法从侧面揭示出企业的发展战略意图。

3 实体关系抽取

依靠人工收集竞争情报的方法不再适用于体量激增的公开信息源，迫切需要智能化的信息收集方法。信息抽取技术应景而生，其目标是抽取文本中的重要信息并结构化存储，以便进一步组织、加工和利用。实体关系抽取是信息抽取中的核心任务，其任务是抽取文本中命名实体之间的关系，例如雇佣关系、家庭关系和合作关系等。实体关系抽取被广泛应用于语义检索、智能问答、机器阅读等自然语言处理领域。命名实体是在第六届消息理解大会（Message Understanding Conference，MUC6）上提出的，其目标是识别文本中的具体实体，例如人名、地点、组织机构、时间、数值等[7]。美国国家标准与技术研究院（National Institute of Standards and Technology，NIST）在 2008 年自动内容抽取（automatic content extraction，ACE）评测方案[8]

中定义了 7 种实体关系类型和 18 种子类型。

O. Etzioni 等 [9] 将实体关系抽取的方法分为 3 类：基于规则的方法、有监督的方法和无监督的方法。

基于规则的方法就是预先定义实体间关系的规则结构，通过模式匹配抽取出实体关系。C. Aone 等 [10] 通过分析文本的内容特征，请领域专家编写实体关系的规则描述，从而抽取实体间的关系。K. Humphreys 等 [11] 首先对文本进行大量的关系规则分析，从中挑选出不会产生错误关系的规则来抽取实体关系。基于规则的方法要求编写规则的人对领域的知识有深入的了解，其优点是抽取速度快，准确率高；缺点是需要人工参与规则的编写，不相关领域难以移植。

有监督的方法是通过机器学习培训人工标注过的语料，训练模型来识别新文本中的实体关系。用于培训语料的机器学习方法有基于核函数方法 [12-13]、逻辑回归方法 [14]、增强解析方法 [15] 以及条件随机场方法 [15-16]，通过这些方法能够以有监督的方式解决关系抽取的问题。该方法的缺点是需要花费大量的时间和精力标注语料，优点是只要有足够的培训语料就能够很好地适应其他领域。

无监督的方法就是通过领域独立的抽取模式自动地发现和标注自己的语料。KnowItAll[17] 是第一个使用无监督方法的信息抽取系统，能够大规模地、领域独立地抽取 Web 页面中的关系，他使用一组通用的模式来自动实例化特定的关系抽取规则。M. Banko 等 [18] 引入了 Open IE 范式，提出开放式关系抽取（open relation extraction，ORE）的概念，开发了 TextRunner 系统，利用启发式规则在宾州树库中自动构建规则，实现大规模的关系抽取。O. Etzioni 等 [19] 根据动词短语识别关系的规则设计了 ReVerb 系统，优先识别关系，然后抽取出该关系的左右实体。L.D. Corro 等 [20] 提出了 ClausIE 系统，用依存语法解析器解析句子，实现关系抽取。无监督的方法优点在于不需要预先定义实体关系的类型或规则，从而方便跨领域的算法移植；缺点在于该方法的抽取结果很依赖于语料库的质量，并且会出现无实际意义的抽取结果。

由于互联网上存在大量的企业并购新闻，需要及时地获取和处理，为了免于人工语义标注和培训语料，同时便于研究领域的移植，本文决定采用无监督的方法抽取企业并购的实体关系。即以动词为主控词，通过 Stanford 依存语法解析抽取中文句子中的动词关系，再找到动词关系的主语实体和宾语实体，实现实体关系抽取。

4 基于实体关系获取企业并购战略情报

商业信息中的企业并购关系描述了企业实体之间由于收购、兼并等活动而形成的相互联系。企业并购关系是一种浅层的知识，通过企业并购关系可以从侧面了解企业的某些事实信息。当将相关的企业并购关系组织起来，形成一个关联网络时，企业并购信息就会转化为一种有用的竞争情报，为企业的战略决策提供支持。本文的主要研究工作就是基于实体关系抽取技术，并以互联网上国内的企业并购关系为例，设计具体的实体关系抽取方案以获取企业的战略情报。

4.1 企业并购关系抽取

企业并购关系的抽取实质是抽取企业实体之间的收购关系。企业并购关系可用关系三元组 <Company1，Verb，Company2> 表示，其中 Company 1 是并购关系的施动者即收购企业，Verb 是并购关系的同义动词，如“兼并”“收购”“购买”等，Company2 是并购关系的受动者，即被收购企业。例如句子“联想宣布以 29 亿美元收购摩托罗拉移动。”可以从中抽取出企业“联想”和“摩托罗拉移动”之间的并购关系 < 联想，收购，摩托罗拉移动 >。本文采用 StanfordParser 实现对中文句子的依存语法解析。StanfordParser 是斯坦福自然语言处理小组提供的依存语法解析工具，能够分析句子的语法结

知识管理论坛，2016(4):266–275
DOI：10.13266/j.issn.2095–5472.2016.032

构，支持英语、中文、德语、法语、阿拉伯语等语言。

将前面的中文例句经过 StanfordParser 解析后产生依存关系有 :nsubj(宣布 -2, 联想 -1), root(ROOT-0, 宣布 -2), prep(收购 -6, 以 -3), nummod(美元 -5, 29 亿 -4), pobj(以 -3, 美元 -5), ccomp(宣布 -2, 收购 -6), nn(移动 -8, 摩托罗拉 -7), dobj(收购 -6, 移动 -8), xsubj(收购 -6, 联想 -1)。为了直观显示句子中各词语之间的依存关系，画出语法依赖图，如图 1 所示：

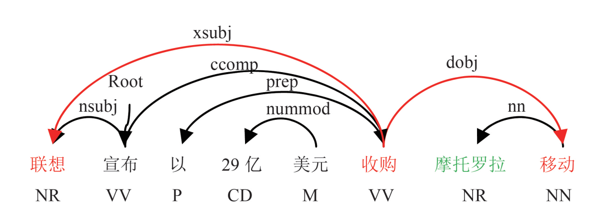


图 1 一条企业并购中文句子的依存解析实例

可以看到主控词动词有两个：“宣布”和“收购”，“宣布”只有一个主语关系“nsubj- 联想”，没有宾语关系; “收购”既有主语关系“xsubj- 联想”，又有宾语关系“dobj- 移动”，所以得到一个关系为 <联想，收购，移动>。但这显然还未完成，“移动”并不等于“摩托罗拉移动”。所以还要考虑名词之间的强修饰关系“nn- 摩托罗拉”，得到最终的关系抽取结果 <联想，收购，摩托罗拉移动>。

4.2 企业属性关系抽取

从企业并购关系 <联想，收购，摩托罗拉移动> 中可以得到一个事实“联想收购了摩托罗拉移动”，但意图是什么？还不是很明了。经过对企业本身的属性研究，企业并购的战略意图可以通过企业所并购公司的经营范围来揭示。例如，摩托罗拉移动的经营范围是智能手机，因此我们可以大概了解到联想收购摩托罗拉移动的战略意图是要近一步加强在智能手机方面的竞争能力。为了进一步挖掘企业并购的战略意图，还需要将被并购企业的经营范围或服务领域找出来，为此本文将被并购企业进行实体链接到百度百科中的词条信息。这样做有两个好处，一是验证所抽取到的被并购企业是否是真实存在的企业实体；二是获取到该企业实体的经营范围属性信息。

百度百科 [21] 是一部内容开放、自由的网络百科全书，旨在创造一个涵盖所有领域知识，服务所有互联网用户的中文知识性百科全书。但是百度百科还没有开放免费的 API，不支持后台直接查询词条信息，只能在前台依靠模式匹配的方法将企业词条信息框中的“经营范围”提取出来，如图 2 所示：



图 2 百度百科的信息框属性

首先，通过编程模拟 URL 查询百科的词条搜索，定位到该企业的详细页面，然后用 HTMLParser 工具抓取该网页到本地并解析成 DOM 树，最后通过遍历该页面的 DOM 树获取到经营范围的值“智能手机”。在实际的抽取当中，有的词条信息框并没有“经营范围”属性，可以考虑近义的信息属性项，如“所属行业”“服务范围”等。



4.3 并购关系抽取技术架构

根据以上的企业并购关系和属性关系的抽取策略，基于垂直搜索和实体关系抽取技术，构建了企业并购关系战略意图揭示系统，其技术架构主要分为数据准备、信息采集、信息处理、信息抽取和意图揭示 5 层，如图 3 所示：

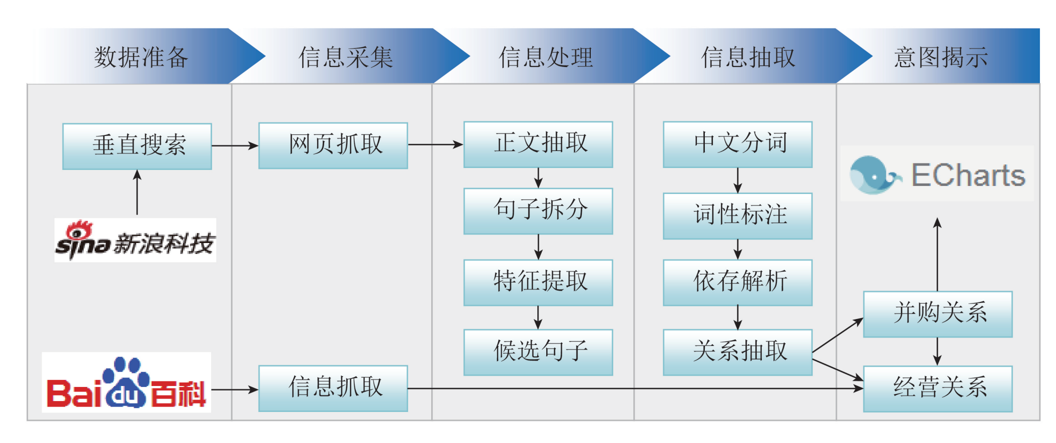


图 3 网络监督情报预警系统技术架构

4.3.1 数据准备

利用百度的网站垂直搜索功能，从新浪科技新闻频道检索与研究主题相关的新闻，再从结果列表中模式匹配出新闻的 URL 以备用。由于后期需要从百度百科中匹配出企业实体的经营领域，该阶段也需要实现对相关词条的检索。

4.3.2 信息采集

用网页抓取工具加载数据准备阶段解析出的相关新闻 URL 列表，提取相关网页的 HTML 文本。

4.3.3 信息处理

主要完成对抓取到的网页进行清洗加工的任务。包括：①正文抽取是抽取文章中的正文内容，过滤掉与新闻报道正文不相干的噪音信息，如超链接、评论等；②句子拆分即将整篇文章拆分成一个个的句子；③特征提取即根据研究主题提取有特征的词语，如并购、兼并、收购等；④候选句子即通过特征词将所需要抽取的主题句子筛选出来。

4.3.4 信息抽取

主要抽取企业实体之间的并购关系。包括：①中文分词即将主题句进行中文分词，便于进行自然语言处理；②词性标注即标注句子中每个词语的词性；③依存解析即通过 StanfordParser 对主题句进行依存句法解析，得到各个词语之间的依存关系；④关系抽取，根据动词关系抽取企业实体之间的并购关系，再根据被并购的企业实体，通过信息采集的信息抓取功能，获取百度百科中的经营属性关系。

4.3.5 意图揭示

通过并购关系和经营关系，为企业提供经过总结的有针对性的竞争情报分析报告，经过 Echarts 可视化工具呈现，达到辅助决策的目的。

5 系统实现

本企业并购情报收集系统框架采用 MVC 设计模式实现三层体系结构：①客户端用户界面使用 Jsp+Echarts 技术，实现战略意图的可视化呈现；②应用服务器端以 Tomcat+Spring Framework 作软件支持平台，使用 java 编程语言实现系统的情报采集、关系抽取和属性挖掘等功能；③数据库服务器端采用 MySQL 5.7，它不仅具有数据仓库功能，而且支持 json，能够实现对关系三元组数据的存储和集成。

5.1 抽取并购关系

实验的数据集采用百度搜索采集，通过百度的网站搜索功能检索新浪科技频道的新闻，检索词式为“阿里巴巴 + 收购”，得到 1200 篇相关报道的网址。用本系统加载网址抓取网

知识管理论坛，2016(4):266–275
DOI：10.13266/j.issn.2095–5472.2016.032

页，抽取正文，提取有“收购”“并购”“兼并”“购买”等动词特征的候选句子 1066 句，然后利用 StanfordParser 解析这些句子从而提取企业并购关系。再从得到的关系集中去除主语不含“阿里”的关系以及含有“或将”“欲”“疑”等词的不确定并购关系。将这些关系的宾语链接到百度百科进行实体识别，并挖掘其所对应的“经营范围或服务范围”属性，形成新的企业属性关系。最后将抽取的并购关系和属性按 json 格式输出成节点和关系数据集，如图 4 所示：

```
Console ☒  @ Javadoc  Problems  Declaration  Search  Call Hierarchy
<terminated> CnExtractor [Java Application] D:\Java\jdk1.7.0_25\bin\javaw.exe (2015年12月
节点:
{category:0, name: '阿里巴巴', value : 5},
{category:1, name: '翰海源', value : 3},
{category:2, name: '计算机软硬件技术开发,咨询,服务,销售', value : 2},
{category:1, name: '365翻译', value : 3},
{category:2, name: '笔译 口译 听译 本地化 定制服务', value : 2},
{category:1, name: '优酷土豆', value : 3},
{category:2, name: '网络视频', value : 2},
{category:1, name: '高德', value : 3},
{category:2, name: '向移动互联网转型', value : 2},
{category:1, name: '优视科技', value : 3},
{category:2, name: '', value : 2},
{category:1, name: '酷盘', value : 3},
{category:2, name: '', value : 2},
{category:1, name: '雅虎', value : 3},
```

```
Console ☒  @ Javadoc  Problems  Declaration  Search  Call Hierarchy
<terminated> CnExtractor [Java Application] D:\Java\jdk1.7.0_25\bin\javaw.exe (2015年12月7日 下午4:23:0
{category:2, name: '软件', value : 2},
关系:
{source : '阿里巴巴', target : '翰海源', weight : 1},
{source : '翰海源', target : '计算机软硬件技术开发,咨询,服务,销售', weight : 1},
{source : '阿里巴巴', target : '365翻译', weight : 1},
{source : '365翻译', target : '笔译 口译 听译 本地化 定制服务', weight : 1},
{source : '阿里巴巴', target : '优酷土豆', weight : 1},
{source : '优酷土豆', target : '网络视频', weight : 1},
{source : '阿里巴巴', target : '高德', weight : 1},
{source : '高德', target : '向移动互联网转型', weight : 1},
{source : '阿里巴巴', target : '优视科技', weight : 1},
{source : '优视科技', target : '', weight : 1},
{source : '阿里巴巴', target : '酷盘', weight : 1},
{source : '酷盘', target : '', weight : 1},
{source : '阿里巴巴', target : '雅虎', weight : 1},
{source : '雅虎', target : '因特网导航服务的网站', weight : 1},
{source : '阿里巴巴', target : '口碑网', weight : 1},
```

图 4 企业并购实体关系抽取结果

从图 4 中可以看到有空值，这是因为从百度百科中挖掘不出相应的经营范围属性值，这时需要人工添加补足。同时还要处理很多重复的关系，最后余下 24 条企业并购关系和 24 条属性关系。

根据互联网行业分类将“经营范围”进一步规范为相应的行业领域，得到被并购企业的行业领域情况，如表 1 所示：



知识管理论坛
2016 年 第 4 期（总第 4 期）

表 1 被并购企业所属领域统计

并购企业	经营范围	所属领域	并购数目（个）
翰海源	计算机软硬件技术开发，咨询，服务，销售	软件开发	7
优视科技	移动互联网软件技术及应用服务提供商	软件开发	
友盟	移动开发者服务平台	软件开发	
擎天科技	软件	软件开发	
粤科	票务电影行业软件	软件开发	
Auctiva	营销管理工具软件	软件开发	
PHOWind	通用型论坛程序	软件开发	
优酷土豆	网络视频	娱乐传媒	5
文化中国	文化传媒	娱乐传媒	
虾米网	与众不同的音乐分享平台	娱乐传媒	
恒大足球	职业足球	娱乐传媒	
南华早报	英文报纸	娱乐传媒	
口碑网	评论分享、消费指南	信息聚合	4
新浪微博	网络信息服务	信息聚合	
Zulily	母婴用品团购	信息聚合	
CNZZ	中国互联网最影响力流量统计网站	信息聚合	
酷盘	网盘文件管理及备份	云计算	2
Vendio	SaaS（软件即服务）提供商	云计算	
雅虎	因特网导航服务的网站	信息搜索	1
天宇朗通	手机专业销售服务	手机通讯	1
一达通	物流企业服务	物流	1
高德	移动地图导航	地图导航	1
365 翻译	笔译口译听译本地化定制服务	教育服务	1
天弘基金	基金 / 证券 / 期货 / 投资	金融	1

5.2 战略意图揭示

本系统实现了根据关系抽取和属性挖掘处理后的结果生成可视化报告，以便进行战略意图分析。如图 5 能够很好地揭示阿里巴巴收购企业公司的意图。其中，粉色的球点为被收购的企业，绿色球点为被收购企业的经营范围，红色球点为被收购企业的行业领域。可以看出，软件开发、娱乐传媒和信息聚合是阿里巴巴收购的重点行业领域，其次是云计算、信息搜索、教育服务、金融等行业。

再将阿里巴巴所并购企业的行业领域用饼图呈现，得到如图 6 所示的企业并购领域分析饼图：

知识管理论坛，2016(4):266-275
DOI：10.13266/j.issn.2095-5472.2016.032

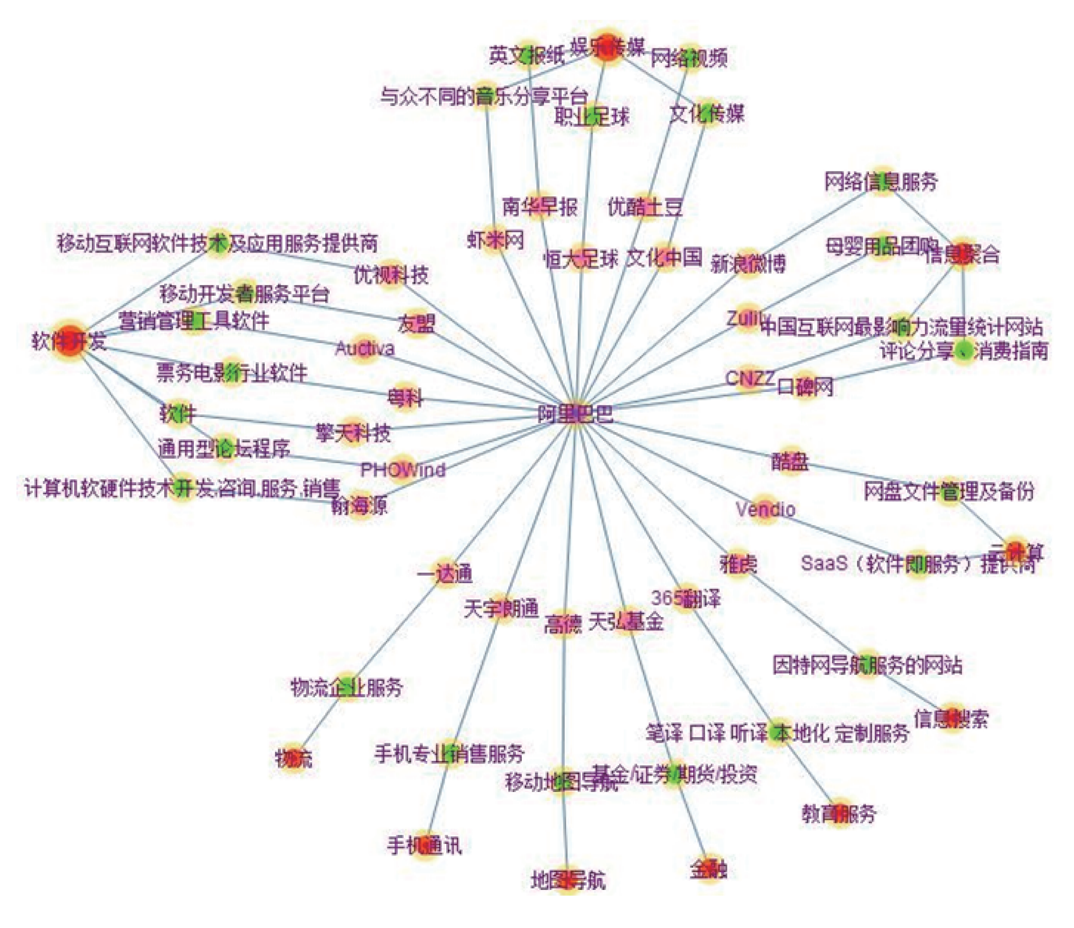


图 5 企业并购战略意图揭示

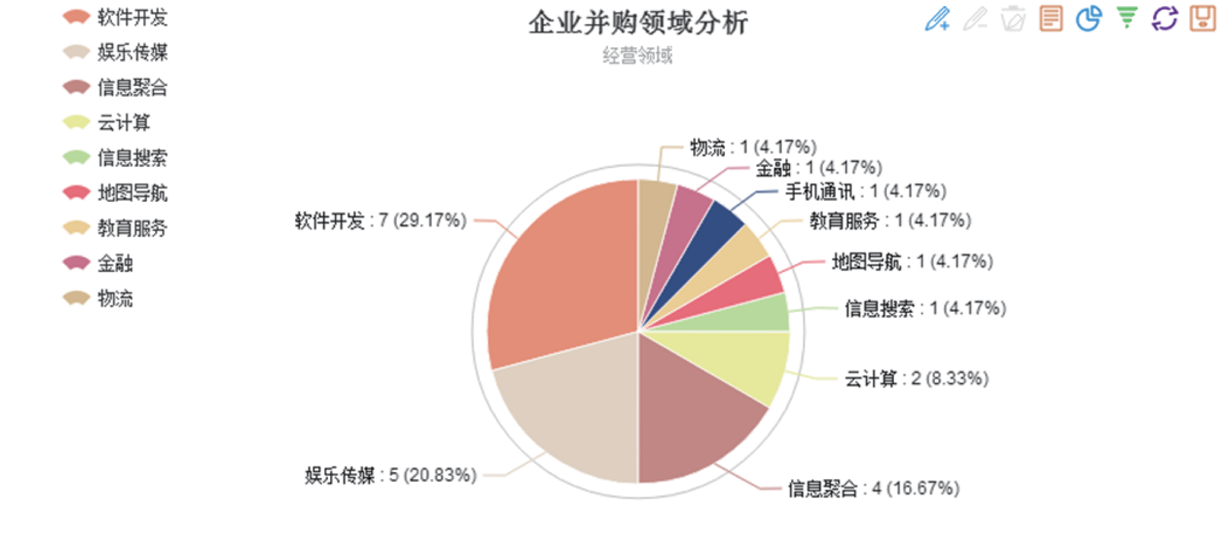


图 6 企业并购领域分析

可以看出，阿里巴巴的战略意图大体可以分为 3 个目标：

（1）构建多元化电商业务生态体系。从并购企业的所属行业领域来看，为了实现成为全世界最大电子商务服务提供商的整体战略愿景，阿里巴巴并购了其核心业务发展所需要的 10 个相关领域，包括软件开发、娱乐传媒、信息聚合、物流及金融等，意在将阿里巴巴的影响延伸到每个用户。通过多元化的并购，既可以进行技术整合，又能够通过垄断市场入口建立一个多元化发展的电商业务生态体系。

（2）打造大数据分析云平台服务能力。在大数据战略上，阿里巴巴收购了 7 家能够进行大数据分析的软件开发公司和 2 家云计算公司，以数据为中心，通过对用户商业核心数据和行为数据的挖掘，找到战略投资方向，进而形成涵盖用户日常生活、商业、社交、学习等领域的数据收集和汇总的生态体系，最终以阿里云平台为承接平台，建立起围绕用户多方面需求的电商业务生态圈。

（3）跨界布局 O2O 商业模式。阿里巴巴转变其纯电商的思维方式，通过跨界收购优酷土豆视频网站、信息聚合和娱乐传媒等 O2O 平台，既可以吸纳互联网用户流量，也可以收集丰富的用户行为数据资源，同时，利用大数据挖掘用户行为数据可以为投放商业广告奠定基础。O2O 即 Online to Offline，是指让互联网成为线下交易的平台。阿里巴巴发展 O2O 商业模式一方面可以辅助支撑电商平台，另一方面更是为了拓展全新的盈利模式，保证其核心竞争力，为企业的竞争情报收集方法提供借鉴。该方法可以解决以往情报服务中只注重实体术语，而忽视实体间关系的不足，将竞争情报分析从传统的自动分类、主题挖掘等分析方法深入到语义挖掘层面，围绕实体之间的关系自动获取战略情报信息，并提供战略意图的可视化呈现，方便企业情报部门快速获取竞争情报。

6 结语

随着企业竞争的加剧，为了寻求长期发展，多元化经营成为企业可持续发展的一个战略。阿里巴巴集团作为一家日益壮大的电子商务公司，通过企业并购发展战略，实现多元化的互联网 O2O 商业模式。本文通过从海量信息中自动获取、挖掘阿里巴巴收购的情报信息，揭示其进行企业并购的战略意图，为我国企业的竞争情报收集方法提供借鉴。该方法可以解决以往情报服务中只注重实体术语，而忽视实体间关系的不足，将竞争情报分析从传统的自动分类、主题挖掘等分析方法深入到语义挖掘层面，围绕实体之间的关系自动获取战略情报信息，并提供战略意图的可视化呈现，方便企业情报部门快速获取竞争情报。

当然该系统方案不可避免地存在需要进一步完善的方面。例如，企业实体及其属性的抽取精度有待进一步提高，由于其表达方式多样、规律性差，还需提出更好的解决方案。可以增加对于时间和收购金额的抽取，丰富企业并购的语义信息。还可以建立完善的领域知识库，如规范的企业机构及其属性信息，能为实体的一致化处理提供完备的信息。

参考文献：

[1] 李国秋 , 吕斌 , 吴鲲 . 企业并购目标选择的竞争情报评估模型——兼论中信证券并购广发证券失败的竞争情报缺失 [J]. 情报理论与实践 ,2008(1):55-59.

[2] NORLING P M, HERRING J P, ROSENKRANS W A, et al. Putting competitive technology intelligence to work[J]. Research-technology management, 2000, 43(5): 23-28.

[3] WAMBUGU P W. An analysis of the effects of competitive intelligence practices on the performance of pharmaceutical companies in Nairobi[J]. International academic journal of human resource and business administration, 2015, 1(4): 1-13.

[4] BRUNER R F. Does M&A pay? A survey of evidence for the decision-maker[J]. Journal of applied finance, 2002, 12(1): 48-68.

[5] 包昌火 , 谢新洲 . 企业竞争情报系统 [M]. 北京 : 华夏出版社 , 2002.

[6] HAMEL G, PRAHALAD C K. Strategic intent[J].Harvard business review, 2005, 83(7): 148-161.

[7] MANSOURI A, AFFENDEY L S, MAMAT A. Named entity recognition approaches[J]. International Journal of computer science and network security, 2008, 8(2): 339-344.

[8] Automatic Content Extraction 2008 Evaluation Plan (ACE08) [EB/OL].[2015-11-15].http://itl.nist.gov/iad/mig/tests/ace/2008/doc/ace08-evalplan.v1.2d.pdf.

[9] ETZIONI O, BANKO M, SODERLAND S, et al.Open information extraction from the web[J]. Communications of the ACM, 2008, 51(12): 68-74.

[10] AONE C, HALVERSON L, HAMPTON T, et al.SRA: description of the IE2 system used for MUC-7[C]//

知识管理论坛，2016(4):266–275
DOI：10.13266/j.issn.2095–5472.2016.032

Proceedings of the seventh message understanding conference(MUC-7).Fairfax:VA,1998.

[11] HUMPHREYS K, GAIZAUSKAS R, AZZAM S,et al.University of Sheffield: description of the LaSIE-II system as used for MUC-7[C]//Proceedings of the seventh message understanding conferences (MUC-7) . San Francisco: Morgan Kaufmann Publishers Inc., 1998.

[12] ZHAO S, GRISHMAN R. Extracting relations with integrated information using kernel methods[C]// Proceedings of the 43rd annual meeting on association for computational linguistics. Morristown:Association for Computational Linguistics, 2005: 419-426.

[13] BUNESCU R, MOONEY R. Subsequence kernels for relation extraction[C]//Advances in neural information processing systems. Cambridge: MIT Press,2005: 171-178.

[14] KAMBHATLA N. Combining lexical, syntactic, and semantic features with maximum entropy models for extracting relations[C]//Proceedings of the ACL 2004 on Interactive poster and demonstration sessions. Stroudsburg: Association for Computational Linguistics, 2004: 22.

[15] MILLER S, FOX H, RAMSHAW L, et al. A novel use of statistical parsing to extract information from text[C]// Proceedings of the 1st North American chapter of the association for computational linguistics conference.San Francisco:Morgan Kaufmann Publishers Inc.,2000: 226-233.

[16] CULOTTA A, MCCALLUM A, BETZ J. Integrating probabilistic extraction models and data mining to discover relations and patterns in text[C]// Proceedings of the main conference on Human Language technology conference of the North American chapter of the association of computational linguistics. New York: ACM, 2006: 296-303.

[17] ETZIONI O, CAFARELLA M, DOWNEY D,et al.Unsupervised named-entity extraction from the web: an experimental study[J]. Artificial intelligence, 2005, 165(1): 91-134.

[18] BANKO M, CAFARELLA M J, SODERL S, et al.Open information extraction from the Web[C]//Proceedings of the 20th international joint conference on artificial intelligence. San Francisco: Morgan Kaufmann Publishers Inc.,2007:2670-2676.

[19] ETZIONI O, FADER A, CHRISTENSEN J, et al. Open information extraction: the second generation[C]// Proceedings of the twenty-second international joint conference on artificial intelligence. Menlo Park: AAAI Press,2011:3-10.

[20] DEL CORRO L, GEMULLA R.ClausIE: clause-based open information extraction[C]//Proceedings of the 22nd international conference on World Wide Web. New York: ACM, 2013: 355-366.

[21] 百度百科 [EB/OL].[2015-11-15].http://baike.baidu.com/view/1.htm.

Strategic Intent Research of Mergers and Acquisitions Based on Entity Relation

Liao Jianjun[1,2,3]
[1]National Science Library, Chinese Academy of Sciences, Beijing 100190
[2]University of Chinese Academy of Sciences, Beijing 100049
[3]Zhengzhou Institute of Aeronautical Industry Management, Zhengzhou 450015

Abstract: [Purpose/significance] This paper proposes an automatic and efficient mining process of the strategic intent for the acquisition issues of enterprise strategic intelligence in competitive intelligence, and guides the enterprises to gather strategic intelligence information and improve the effectiveness of the strategic intent mining. **[Method/process]** This paper explored the system solution of enterprise strategic intelligence acquisition by introducing the technology of entity relation extraction. Then it took enterprise merger events as examples, applied the relation extraction technology to collect the enterprise merger and acquisition relation and revealed the strategic intent of enterprise mergers and acquisitions by entity attribute extraction, and verified the effectiveness of the extraction system through the merger and acquisition cases of Alibaba Group. **[Result/conclusion]** This solution can efficiently and automatically extract the enterprise merger and acquisition relation from massive web information and mine enterprises business areas. Through the visual technique to reveal the strategic intents of enterprise mergers and acquisitions, the purposes of automatic and fast acquisition of enterprise competitive intelligence are achieved.

Keywords: entity relation relation extraction mergers and acquisitions strategic intent competitive intelligence